

Risk aggregation and stochastic claims reserving in disability insurance

Boualem Djehiche* Björn Löfdahl †

August 27, 2014

Abstract

We consider a large, homogeneous portfolio of life or disability annuity policies. The policies are assumed to be independent conditional on an external stochastic process representing the economic-demographic environment. Using a conditional law of large numbers, we establish the connection between claims reserving and risk aggregation for large portfolios. Further, we derive a partial differential equation for moments of present values. Moreover, we show how statistical multi-factor intensity models can be approximated by one-factor models, which allows for solving the PDEs very efficiently. Finally, we give a numerical example where moments of present values of disability annuities are computed using finite-difference methods and Monte Carlo simulations.

Keywords: Disability insurance, stochastic intensities, conditional independence, risk aggregation, stochastic claims reserving, mimicking.

1 Introduction

The upcoming Solvency II regulatory framework brings many new challenges to the insurance industry. In particular, the new regulations suggest a new mindset regarding the valuation and risk management of insurance products. Historically, premiums and reserves are calculated under the assumption that the underlying transition intensities of death, disability onset, recovery and so on are deterministic. While the estimations should be prudent, this still implies that the systematic risk, i.e. the risk arising from uncertainty of the future development of the hazard rates, is not taken into account. This may have an impact on pricing as well as capital charges. In the Solvency II standard model, capital charges are computed using a scenario based approach, and the capital charge is given as the difference between the present value under best estimate assumptions, and the present value in a certain shock scenario. As an alternative, insurers may adopt an internal model, which should be based on a Value-at-Risk approach.

Facing these challenges, a plethora of stochastic intensity models have appeared, in particular for modelling mortality. However, these works have largely focused on either calibration or on pricing a single policy under a suitable market implied measure. The risk management aspect has been left largely untouched, although there are notable

*Department of Mathematics, KTH Royal Institute of Technology, Sweden, boualem@kth.se.

†Department of Mathematics, KTH Royal Institute of Technology, Sweden, bjornlg@kth.se.

exceptions. Dahl [8] derives a pricing PDE for a wide class of life derivatives under a one-factor stochastic intensity model. Dahl points out that shocks from a one-factor model affect all cohorts equally, and that a multi-factor model across cohorts might be more realistic, although it would not offer any further insights. Dahl and Møller [9] consider pricing and hedging of life insurance liabilities with systematic mortality risk. Biffis [4] considers annuities pricing under affine mortality models. Ludkovski and Young [18] consider indifference pricing under stochastic hazard. Norberg [19] derives an ODE for moments of present values assuming deterministic hazard rates.

While stochastic mortality models have been thoroughly studied in the literature, stochastic disability models have not received the same attention. Levantesi and Menzietti [17] consider stochastic disability and mortality in the Solvency II context. The approach covers both systematic and idiosyncratic risk, and is suitable for small portfolios. Christiansen *et al.* [7] suggest an internal model for Solvency II based on the forecasting technique of Hyndman and Ullah [13]. The approach includes fitting an intensity model over a range of time periods, and fitting a time series model to the time series of parameter estimates. The future development of the intensities is obtained by forecasting or simulation of the time series model.

In this paper, we consider a large, homogeneous portfolio of life or disability annuity policies. The policies are assumed to be independent conditional on an external stochastic process representing the economic-demographic environment. Using a conditional law of large numbers, we show that the aggregated annuity cash flows can be approximated by its conditional expectation, an expression much akin to the actuarial reserve formula. This result highlights the connection between risk aggregation and claims reserving for large portfolios. Further, we derive a partial differential equation for moments of these present values. Moreover, we consider statistical multi-factor intensity models, and suggest methods for reducing their dimensionality. Using the so-called mimicking technique introduced by Krylov [14], we suggest approximating multi-factor models by one-factor models, which allows for solving the PDEs very efficiently. Finally, we give a numerical example where moments of present values of disability annuities are computed using finite-difference methods and Monte Carlo simulations.

The paper is organized as follows. In Section 2, consider an annuity policy under a simple stochastic intensity model. We derive a PDE for computing moments of the random present value of such policies. In Section 3, we examine the aggregated cash flows from a large, homogeneous portfolio of insurance policies, and highlight the connection between risk aggregation and claims reserving. In Section 4, we consider the specific application of disability insurance, and show how a class of statistical models can be incorporated into the pricing PDEs. In Section 5, we present numerical results based on disability data from the Swedish insurance company Folksam.

2 Stochastic claims reserving

Let τ^1, τ^2, \dots be random event times (e.g. times of death or recovery from disability), and let

$$N_t^k = I\{\tau^k \leq t\}, \quad k \geq 1. \quad (1)$$

Further, define the processes

$$N^k = (N_t^k)_{t \geq 0}, \quad k \geq 1, \quad (2)$$

and let

$$\mathcal{F}^N = (\mathcal{F}_t^N)_{t \geq 0} = (\mathcal{F}_t^{N^1} \vee \mathcal{F}_t^{N^2} \vee \dots)_{t \geq 0} \quad (3)$$

denote the filtration generated by N^1, N^2, \dots . Now, let Z be a stochastic process with natural filtration $\mathcal{F}^Z = (\mathcal{F}_t^Z)_{t \geq 0}$. Here, N_t^k denotes the state of an insured individual at time t , τ^k represents the corresponding death or recovery time, and Z_t represents the state of the economic-demographic environment. We assume that N^1, N^2, \dots are independent conditional on \mathcal{F}_∞^Z , and that the $\mathcal{F}^Z \vee \mathcal{F}^N$ -intensity of N^k is the process λ^k of the form

$$\lambda_t^k = q(t, Z_t)(1 - N_t^k), \quad t \geq 0. \quad (4)$$

Consider an annuity policy paying $g(t, Z_t)$ continuously as long as $N_t^k = 0$, until a fixed future time T . This type of annuity allows for payments from the contract to depend on time as well as the state of the economic-demographic environment. For example, the contract could be inflation-linked and contain a deferred period. The random present value L_t^k of this policy can be written as

$$L_t^k = \int_t^T g(s, Z_s)(1 - N_s^k) e^{-\int_t^s r(u) du} ds, \quad (5)$$

where the short rate r is assumed to be adapted to \mathcal{F}^Z . Further, the time t reserve for this contract is given by $E[L_t^k | \mathcal{F}_t^Z \vee \mathcal{F}_t^N]$, that is, the expected value given the history of the policy and of the environment. Our goal is to find an efficient way to compute this reserve. First, we need the following result, which is given in a slightly different form in Norberg's concise introduction to stochastic intensity models [20].

Proposition 1 *Assume that $E[|\lambda_t^k|] < \infty$ for each $k, t \geq 0$. Then, for $s \geq t$,*

$$E[1 - N_s^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] = P(\tau^k > s | \mathcal{F}_s^Z \vee \mathcal{F}_t^N) = (1 - N_t^k) e^{-\int_t^s q(u, Z_u) du}. \quad (6)$$

Proof. First, note that the process $(M_s^k)_{s \geq 0}$ defined by

$$M_s^k = N_s^k - \int_0^s \lambda_u^k du \quad (7)$$

is a $\mathcal{F}^Z \vee \mathcal{F}^N$ -martingale [20, p. 106]. For $s \geq t$, let $Y_s^k = P(\tau^k > s | \mathcal{F}_s^Z \vee \mathcal{F}_t^N) = E[1 - N_s^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N]$. Using (7), we have

$$\begin{aligned} Y_s^k &= E[1 - N_s^k + \int_0^s \lambda_u^k du - \int_0^s \lambda_u^k du | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] \\ &= 1 - N_t^k + \int_0^t \lambda_u^k du - E[\int_0^s \lambda_u^k du | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] \\ &= 1 - N_t^k + \int_0^t \lambda_u^k du - \int_0^t \lambda_u^k du - E[\int_t^s \lambda_u^k du | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] \\ &= 1 - N_t^k - \int_t^s q(u, Z_u) E[1 - N_u^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] du \\ &= 1 - N_t^k - \int_t^s q(u, Z_u) E[1 - N_u^k | \mathcal{F}_u^Z \vee \mathcal{F}_t^N] du \\ &= 1 - N_t^k - \int_t^s q(u, Z_u) Y_u^k du. \end{aligned} \quad (8)$$

Differentiating the above expression, we obtain

$$\begin{cases} dY_s^k = -q(s, Z_s)Y_s^k ds, & s > t, \\ Y_t^k = 1 - N_t^k, \end{cases} \quad (9)$$

with solution $Y_s^k = (1 - N_t^k)e^{-\int_t^s q(u, Z_u)du}$. \square

Using Proposition 1, we immediately obtain

$$\begin{aligned} E[L_t^k | \mathcal{F}_t^Z \vee \mathcal{F}_t^N] &= E[E[L_t^k | \mathcal{F}_T^Z \vee \mathcal{F}_t^N] | \mathcal{F}_t^Z \vee \mathcal{F}_t^N] \\ &= E[E[\int_t^T g(s, Z_s)(1 - N_s^k)e^{-\int_t^s r(u)du} ds | \mathcal{F}_T^Z \vee \mathcal{F}_t^N] | \mathcal{F}_t^Z \vee \mathcal{F}_t^N] \\ &= (1 - N_t^k)E[\int_t^T g(s, Z_s)e^{-\int_t^s q(u, Z_u)du} e^{-\int_t^s r(u)du} ds | \mathcal{F}_t^Z \vee \mathcal{F}_t^N]. \end{aligned} \quad (10)$$

Note that if the environment process Z is replaced by a deterministic function, the functional V_t defined by

$$V_t = \int_t^T g(s, Z_s)e^{-\int_t^s q(u, Z_u)du} e^{-\int_t^s r(u)du} ds \quad (11)$$

corresponds to the time t -reserve of a policy paying g monetary units continuously. Now, since q and g are functions of the stochastic process Z , V_t is a random variable, and the reserve depends on the distribution of V_t . In the case where Z is a Markov process, a natural candidate for the time t reserve of an active contract is the function $v(t, z)$ given by

$$v(t, z) = E[V_t | Z_t = z] = E^{t, z} \left[\int_t^T g(s, Z_s)e^{-\int_t^s q(u, Z_u)du} e^{-\int_t^s r(u)du} ds \right]. \quad (12)$$

Let $\bar{q}(t, z) = q(t, z) + r(t)$, and assume that \bar{q} is lower bounded, g is continuous and bounded, and that Z is a Markov process with infinitesimal generator \mathcal{A} . Then, $v(t, z)$ given by (12) is a Feynman-Kac functional, satisfying the backward PDE

$$\begin{cases} -\frac{\partial v}{\partial s} + \bar{q}(s, z)v = \mathcal{A}v + g(s, z), & t \leq s < T, \\ v(T, z) = 0. \end{cases} \quad (13)$$

For risk management purposes, it is not enough to be able to compute expected values. Often, it is necessary to estimate moments or quantiles. Moments of V_t can be found using the following result.

Proposition 2 *Let $\bar{q} = q + r$, and assume that \bar{q} is lower bounded, g is continuous and bounded, and that Z is a Markov process with generator \mathcal{A} . Then, for $n \geq 1$, $v_n(t, z) = E^{t, z}[V_t^n]$ satisfies the backward PDE*

$$\begin{cases} -\frac{\partial v_n}{\partial s} + n\bar{q}(s, z)v_n = \mathcal{A}v_n + ng(s, z)v_{n-1}, & t \leq s < T, \\ v_n(T, z) = 0, \end{cases} \quad (14)$$

where, naturally, $v_0(t, z) = E^{t, z}[V_t^0] = 1$.

Proof. Differentiating V_t , we obtain

$$dV_t = (\bar{q}_t V_t - g_t) dt. \quad (15)$$

Therefore,

$$d(V_t^n) = nV_t^{n-1}(\bar{q}_t V_t - g_t) dt = (n\bar{q}_t V_t^n - ng_t V_t^{n-1}) dt. \quad (16)$$

Multiplying with the integrating factor $e^{-\int_0^t n\bar{q}_u du}$, integrating and using $V_T^n = 0$, we have

$$V_t^n = \int_t^T ng(s, Z_s) V_s^{n-1} e^{-\int_t^s n\bar{q}_u du} ds. \quad (17)$$

Taking conditional expectations and using the Markov property of Z ,

$$\begin{aligned} E^{t,z}[V_t^n] &= E^{t,z}\left[\int_t^T ng(s, Z_s) V_s^{n-1} e^{-\int_t^s n\bar{q}_u du} ds\right] \\ &= E^{t,z}\left[\int_t^T E[ng(s, Z_s) V_s^{n-1} e^{-\int_t^s n\bar{q}_u du} | \mathcal{F}_s^Z] ds\right] \\ &= E^{t,z}\left[\int_t^T ng(s, Z_s) E[V_s^{n-1} | Z_s] e^{-\int_t^s n\bar{q}_u du} ds\right] \\ &= E^{t,z}\left[\int_t^T ng(s, Z_s) v_{n-1}(s, Z_s) e^{-\int_t^s n\bar{q}_u du} ds\right]. \end{aligned} \quad (18)$$

From the Feynman-Kac formula, it follows immediately that $v_n(t, z)$ satisfies the PDE (14), see e.g. Friedman [10, Theorem 5.3] for details. \square

Proposition 2 can be used to find the k 'th moment of V_t by solving the PDE (14) for $n = 1, \dots, k$ iteratively. This is useful since it is often faster to numerically solve a PDE than to perform a Monte Carlo simulation, especially for this type of path-dependent problem.

3 Risk aggregation

We now consider the risk aggregation problem. For a portfolio consisting of annuity policies for the population N^1, N^2, \dots, N^n , the random present value $L_t^{(n)}$ becomes

$$L_t^{(n)} = \sum_{k=1}^n L_t^k = \sum_{k=1}^n \int_t^T g(s, Z_s) (1 - N_s^k) e^{-\int_t^s r(u) du} ds. \quad (19)$$

We will now investigate the properties of $L^{(n)}$ as the number of policies grows large.

Proposition 3 *Conditional on $\mathcal{F}_T^Z \vee \mathcal{F}_t^N$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} L_t^{(n)} - \frac{1}{n} \sum_{k=1}^n (1 - N_t^k) V_t = 0 \quad a.s., \quad (20)$$

where V_t is given by (11).

Proof. Since N_s^1, N_s^2, \dots are independent conditional on $\mathcal{F}_s^Z \vee \mathcal{F}_t^N$ with

$$\sum_{k=1}^{\infty} \frac{E[(N_s^k - E[N_s^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N])^2 | \mathcal{F}_s^Z \vee \mathcal{F}_t^N]}{k^2} \leq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty, \quad (21)$$

it follows from the conditional Law of Large Numbers (see Prakasa Rao [22, Theorem 6]) that, conditional on $\mathcal{F}_s^Z \vee \mathcal{F}_t^N$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n N_s^k - E\left[\frac{1}{n} \sum_{k=1}^n N_s^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N\right] = 0 \quad a.s. \quad (22)$$

This implies that, conditional on $\mathcal{F}_T^Z \vee \mathcal{F}_t^N$,

$$\begin{aligned} & \frac{1}{n} L_t^{(n)} - E\left[\frac{1}{n} L_t^{(n)} | \mathcal{F}_T^Z \vee \mathcal{F}_t^N\right] \\ &= \frac{1}{n} \sum_{k=1}^n \int_t^T g(s, Z_s) (1 - N_s^k) e^{-\int_t^s r(u) du} ds \\ & \quad - E\left[\frac{1}{n} \sum_{k=1}^n \int_t^T g(s, Z_s) (1 - N_s^k) e^{-\int_t^s r(u) du} ds | \mathcal{F}_T^Z \vee \mathcal{F}_t^N\right] \\ &= \frac{1}{n} \sum_{k=1}^n \int_t^T g(s, Z_s) (E[N_s^k | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] - N_s^k) e^{-\int_t^s r(u) du} ds \rightarrow 0 \quad a.s., \end{aligned} \quad (23)$$

by (22) and the conditional dominated convergence theorem. Now, using Proposition 1, we have

$$\begin{aligned} E\left[\frac{1}{n} L_t^{(n)} | \mathcal{F}_T^Z \vee \mathcal{F}_t^N\right] &= \int_t^T \frac{1}{n} \sum_{k=1}^n E[g(s, Z_s) (1 - N_s^k) | \mathcal{F}_s^Z \vee \mathcal{F}_t^N] e^{-\int_t^s r(u) du} ds \\ &= \int_t^T \frac{1}{n} \sum_{k=1}^n (1 - N_t^k) g(s, Z_s) e^{-\int_t^s q(u, Z_u) du} e^{-\int_t^s r(u) du} ds \\ &= \frac{1}{n} \sum_{k=1}^n (1 - N_t^k) \int_t^T g(s, Z_s) e^{-\int_t^s q(u, Z_u) du} e^{-\int_t^s r(u) du} ds \\ &= \frac{1}{n} \sum_{k=1}^n (1 - N_t^k) V_t. \end{aligned} \quad (24)$$

The claim follows from (23) and (24). \square

When the portfolio is large enough, Proposition 3 motivates the approximation

$$L_t^{(n)} \approx \sum_{k=1}^n (1 - N_t^k) V_t. \quad (25)$$

Hence, in order to determine the distribution of the present value of the portfolio given the history of the environment and the policies, it suffices to consider the random variable V_t . Indeed, all the individual risks are diversified away, and only the systematic risk, that is, the risk that the economic-demographic environment changes, remains.

This is formalized through the random variable V_t . In particular, an approximate p -quantile of the random present value of the portfolio is given by the relation

$$F_{L_t^{(n)}}^{-1}(p) \approx F_{\sum_{k=1}^n (1-N_t^k)V_t}^{-1}(p) = \sum_{k=1}^n (1-N_t^k)F_{V_t}^{-1}(p), \quad (26)$$

where the equality follows from the positive homogeneity of the quantile function. This result is analogous to the loan portfolio risk result of Vasicek [24], which is the foundation of the Basel regulatory credit risk framework. In the Basel framework, the homogeneity requirement of the portfolio is relaxed to allow for efficient approximation of portfolio Value-at-Risk and capital allocation, which possibly suggests that it can also be considered in this application.

Properties of V_t can be investigated using simulation or PDE techniques. Further, the time t reserve for the entire portfolio is given by

$$E[L_t^{(n)}|\mathcal{F}_t^Z \vee \mathcal{F}_t^N] = E[E[L_t^{(n)}|\mathcal{F}_T^Z \vee \mathcal{F}_t^N]|\mathcal{F}_t^Z \vee \mathcal{F}_t^N] = \sum_{k=1}^n (1-N_t^k)E[V_t|\mathcal{F}_t^Z], \quad (27)$$

and the amount of money allocated to each active policy at time t is simply $E[V_t|\mathcal{F}_t^Z]$. Based on these considerations, the problem of risk aggregation is closely connected to the problem of claims reserving.

We conclude this section with some comments regarding the Solvency II framework. In the Solvency II standard model, capital charges are computed using a scenario based approach, and the capital charge is given as the difference between the present value under best estimate assumptions, and the present value in a certain shock scenario. As an alternative, insurers may adopt an internal model, which should be based on a Value-at-Risk approach over a one-year time horizon. For instance, the capital charge may be taken to be the Economic Capital, i.e. the difference between the time t value and the p -quantile of the value at time $t + 1$. We stress the fact that the approximate portfolio quantile given by (26) represents the risk over the entire policy period, i.e. it can be used to compute Value-at-Risk over $T - t$ years. Thus, a topic for future research would be to find an extension of the above result, compatible with the Solvency II framework.

4 Application to disability insurance

In this section, we consider an example from disability insurance. We seek to compute moments of V_t for which the process Z , representing the economic-demographic environment, is constructed from a generalized linear model for disability recovery probabilities. For simplicity, we will assume that the short rate is deterministic. As we will see below, Z is typically non-Markov, and we cannot directly use the Feynman-Kac formula to compute moments of V_t . We will consider two possible solutions to this problem. First, we construct a multivariate Markov process with Z as one of its component. This turns out to work well in some special cases. Second, we will rely on the so-called mimicking technique to obtain a reliable approximation of V_t .

4.1 A stochastic termination model

Following Aro, Djehiche and Löfdahl [1], the probability $p_{\nu_t}(x, d)$ that the disability of an individual with disability inception age x and disability duration d is terminated

within $[d, d + \Delta d)$ is given by

$$p_{\nu_t}(x, d) = \frac{\exp \left\{ \sum_{i=1}^n \phi^i(x) \sum_{j=1}^m \psi^j(d) \nu_t^{i,j} \right\}}{1 + \exp \left\{ \sum_{i=1}^n \phi^i(x) \sum_{j=1}^m \psi^j(d) \nu_t^{i,j} \right\}}, \quad (28)$$

where ϕ and ψ are basis functions in x and d , respectively, and ν is an $n \times m$ -dimensional stochastic process. For simplicity, the termination intensity $q(d, \nu_t)$ is approximated to be piecewise constant over a small time period Δd , i.e. it is given by the relation

$$p_{\nu_t}(x, d) = 1 - \exp \left\{ -q(d, \nu_t) \Delta d \right\}. \quad (29)$$

In the present context, the duration d is simply assumed to be 0 at time $t = 0$. Using this, together with (28)-(29), we obtain, for a fixed x and Δd , the following approximation for the intensity q :

$$q(t, \nu_t) = \frac{1}{\Delta d} \log \left(1 + \exp \left\{ \sum_{i=1}^n \phi^i(x) \sum_{j=1}^m \psi^j(t) \nu_t^{i,j} \right\} \right). \quad (30)$$

Given a suitable stochastic process form for ν , we may solve the PDE (14) with nm space dimensions. However, this is not very efficient when nm is large. To obtain a more tractable model, we will try to reduce the number of dimensions.

4.2 Reducing the dimensionality

Define the process $Z = \{Z_t\}_{t \geq 0}$ by

$$Z_t = \sum_{i=1}^n \phi^i(x) \sum_{j=1}^m \psi^j(t) \nu_t^{i,j}, \quad (31)$$

and define the function f by

$$f(\cdot) = \frac{1}{\Delta d} \log(1 + \exp(\cdot)), \quad (32)$$

so that we have

$$q(t, \nu_t) = f(Z_t), \quad t \geq 0. \quad (33)$$

It is easily seen that we can rewrite Z_t on vector form as

$$Z_t = a(t)^T \nu_t, \quad (34)$$

with

$$a(t)^T = (\phi^1(x) \psi^1(t), \dots, \phi^n(x) \psi^m(t)), \quad (35)$$

$$\nu_t = (\nu_t^{1,1}, \dots, \nu_t^{n,m}). \quad (36)$$

From now on, we restrict our attention to the case where ν can be written as

$$\nu_t = \nu_0 + \mu t + AW_t, \quad (37)$$

where W is an nm -dimensional standard Brownian motion with independent components, $\mu \in \mathbb{R}^{nm}$ and $A \in \mathbb{R}^{nm \times nm}$ is the Cholesky factorization of the covariance matrix Σ of ν . In principle, any dynamic for ν_t is possible. The random walk is a

natural choice, since it is easy to fit and simulate, and has been the model of choice in e.g. Christiansen *et al.* [7]. If a is locally bounded, this modelling choice guarantees that the assumption in Proposition 1 is satisfied, since, in view of (32)-(33), we have

$$E[f(Z_t)] = E\left[\frac{1}{\Delta d} \log(1 + e^{Z_t})\right] \leq \frac{\log 2}{\Delta d} + \frac{1}{\Delta d} E[|Z_t|] < \infty. \quad (38)$$

Next, consider the dynamics of Z . The Itô formula yields, using (37) and (34),

$$dZ_t = (\dot{a}^T \nu_t + a^T \mu)dt + a^T AdW_t, \quad (39)$$

provided that \dot{a} exists. This expression cannot directly be written on the form

$$dZ_t = \alpha(t, Z_t)dt + \gamma(t)dW_t, \quad (40)$$

and therefore it is not a 1-dimensional Itô diffusion. In general, it is not even a Markov process. This is due to the time dependence of a , a property which originates from the fact that the termination intensity depends on the duration of the illness. This property cannot easily be relaxed.

To remedy this, it may be possible to construct a process \hat{Z} of the form (40), identical to Z in law. This would imply

$$V_t = \int_t^T g(s, \nu_s) e^{-\int_t^s q(u, \nu_u) du} e^{-\int_t^s r(u) du} ds \quad (41)$$

$$\stackrel{d}{=} \int_t^T g(s, \hat{Z}_s) e^{-\int_t^s f(\hat{Z}_u) du} e^{-\int_t^s r(u) du} ds =: \hat{V}_t, \quad (42)$$

and, more importantly, that $(V) = (\hat{V})$, i.e. that the processes V and \hat{V} are identical in law. According to Øksendal [21, Theorem 8.4.3], $(\hat{Z}) = (Z)$ if and only if

$$\alpha(t, Z_t) = E[\dot{a}^T \nu_t + a^T \mu | \mathcal{F}_t^Z] \quad (43)$$

$$\gamma^2(t) = a^T A A^T a. \quad (44)$$

Unfortunately, the conditional expectation (43) is in general not easy to compute. We now turn our attention to a special case where it is possible to construct a multivariate Markov process that contains Z .

4.2.1 Construction of a multivariate Markov process

We now consider the case where each component of a is either constant or linear in t . As an example, we take the model from Section 4.1 with basis functions

$$\begin{aligned} \phi^1(x) &= \frac{64-x}{39}, & \phi^2(x) &= \frac{x-25}{39} \\ \psi^1(t) &= 1, & \psi^2(t) &= t. \end{aligned}$$

Aro, Djehiche and Löfdahl [1] fit this model to data from a Swedish insurance company and suggest that it can be seen as a middle ground model when considering goodness of fit versus tractability. Here, it proves to be an interesting special case which allows us to construct a multivariate Markov process from a non-Markovian one.

Consider the vector valued Markov process $Z = (Z^1, Z^2)$ defined by

$$\begin{cases} Z_t^1 &= a^T \nu_t, \\ Z_t^2 &= \dot{a}^T \nu_t. \end{cases} \quad (45)$$

The process Z satisfies the system of stochastic differential equations

$$\begin{cases} dZ_t^1 &= (Z_t^2 + a^T \mu)dt + a^T AdW_t, \\ dZ_t^2 &= \dot{a}^T \mu dt + \dot{a}^T AdW_t. \end{cases} \quad (46)$$

By [21, Theorem 8.4.3], Z is identical in law to the process $\widehat{Z} = (\widehat{Z}^1, \widehat{Z}^2)$ defined by

$$d\widehat{Z}_t = \alpha(t, \widehat{Z}_t)dt + \gamma(t)d\widehat{W}_t, \quad (47)$$

where

$$\alpha(t, Z_t) = E\left[\begin{pmatrix} Z_t^2 + a^T \mu \\ \dot{a}^T \mu \end{pmatrix} \middle| \mathcal{F}_t^Z\right] = \begin{pmatrix} Z_t^2 + a^T \mu \\ \dot{a}^T \mu \end{pmatrix}, \quad (48)$$

$$\gamma(t)\gamma(t)^T = \begin{pmatrix} a^T A \\ \dot{a}^T A \end{pmatrix} \begin{pmatrix} a^T A \\ \dot{a}^T A \end{pmatrix}^T = \begin{pmatrix} a^T \Sigma a & a^T \Sigma \dot{a} \\ a^T \Sigma \dot{a} & \dot{a}^T \Sigma \dot{a} \end{pmatrix}, \quad (49)$$

and \widehat{W} is a two-dimensional standard Wiener process. Thus, we have effectively reduced the process ν to the two-dimensional process \widehat{Z} , and we may compute moments of present values by solving the PDEs (14) with the generator $\widehat{\mathcal{A}}$ of \widehat{Z} and termination intensity $q(t, \widehat{Z}_t) = f(\widehat{Z}_t^1)$.

This recipe can easily be extended to the case where $a^{(k)}$, the k 'th derivative of a w.r.t. time, is constant. Then, the system (46) becomes a system of $k + 1$ SDEs, and the process defined by (47) will have $k + 1$ driving Wiener processes. Thus, if $k + 1 < nm$, that is, if the number of driving Wiener processes is smaller than the number of parameters in the statistical model, the dimensionality of the problem can be reduced, while still preserving all probabilistic properties of the system.

4.2.2 Mimicking the killed environment process

It is not always possible to construct a multivariate Markov process containing Z as above, and even if it is possible, it is not certain that the number of dimensions will be reduced. For example consider the model from Section 4.1 with basis functions

$$\begin{aligned} \phi^1(x) &= \frac{64 - x}{39}, & \phi^2(x) &= \frac{x - 25}{39} \\ \psi^1(t) &= 1, & \psi^2(t) &= e^{-t}, & \psi^3(t) &= e^{-2t}. \end{aligned}$$

It is immediate that we cannot apply the recipe of Section 4.2.1. As an alternative, we will rely on an idea suggested by Krylov [14] to construct a Markov process \widehat{Z} that mimics certain features of the behavior of the process Z such as

$$\widehat{Z}_t \stackrel{d}{=} Z_t, \quad t \geq 0. \quad (50)$$

Proposition 4 below displays a general result about existence of the Markov process \widehat{Z} when Z is a non-Markov diffusion. This result appeared first in Krylov [14] and extended in Gyöngy [11] and Borkar [5] and generalized in various ways to Lévy processes and semimartingales in Bhatt and Borkar [3], Kurtz and Stockbridge [15, 16], Bentata and Cont [2], and Bouhadou and Ouknine [6]. The process \widehat{Z} is often called *Markovian projection* or *mimicking process* of Z .

Proposition 4 (Kurtz and Stockbrigde [15], Corollary 4.3)

When Z satisfies

$$Z_t = Z_0 + \int_0^t \beta(s)ds + \int_0^t \delta(s)dW_s, \quad (51)$$

where, W is an \mathbb{R}^d -valued \mathcal{F}_t -Brownian motion; δ and β are measurable, \mathcal{F}_t -adapted processes taking values in the set of $d \times d$ matrices $\mathbb{M}^{d \times d}$ and \mathbb{R}^d , respectively; and Z_0 is \mathbb{R}^d -valued and \mathcal{F}_0 -measurable. Then there exist measurable functions $\sigma : [0, \infty) \times \mathbb{R}^d \mapsto \mathbb{M}^{d \times d}$ and $b : \mathbb{R}^d \mapsto \mathbb{R}^d$, an \mathbb{R}^d -valued Brownian motion \widehat{W} , and a process \widehat{Z} satisfying

$$\widehat{Z}_t = \widehat{Z}_0 + \int_0^t b(s, \widehat{Z}_s)ds + \int_0^t \sigma(s, \widehat{Z}_s)d\widehat{W}_s, \quad (52)$$

such that for each $t \geq 0$,

$$(Z_t, E[\beta(t)|Z_t], E[\delta(t)\delta^T(t)|Z_t]) \stackrel{d}{=} (\widehat{Z}_t, b(t, \widehat{Z}_t), \sigma(t, \widehat{Z}_t)\sigma^T(t, \widehat{Z}_t)). \quad (53)$$

For the sequel, we set

$$v^Z(t, z) = v(t, z), \quad (54)$$

whenever $Z_t = z$. When the intensity q is a constant, it is immediate that

$$v^{\widehat{Z}}(t, \widehat{Z}_t) \stackrel{d}{=} v^Z(t, Z_t) \quad (55)$$

holds whenever the property (53) remains true. A counter-example constructed by Borkar [5] suggests that it is not always possible to obtain a Markov process \widehat{Z} whose finite dimensional distributions agree with those of the process Z . Therefore, (55) may not hold when the discount factor q depends of Z . Kurtz and Stockbrigde [15, Theorem 5.1] do construct a Markov process \widehat{Z} for which (55) holds, even when the discount factor q depends on Z , but the t -marginal distributions of \widehat{Z} and Z may not be identical i.e. \widehat{Z} does not mimic Z .

A closer look at the t -reserve $v^Z(t, Z_t)$ suggests that we should mimic the process \bar{Z} obtained by 'killing' Z at rate q in the sense described e.g. in Rogers and Williams [23, Section III.18]. The intuitive idea behind killing is that \bar{Z} agrees with Z up to time $\bar{\tau}$ and $\bar{Z}_t = \partial$, $t \geq \bar{\tau}$, where ∂ is some absorbing state, and

$$\bar{P}(\bar{\tau} > t | \mathcal{F}_t^Z) = e^{-\int_0^t q(s, Z_s)ds}. \quad (56)$$

Given a process Z on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, the process \bar{Z} obtained by 'killing' Z at rate q is defined on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathcal{F}}_t, \bar{P})$ by

$$\bar{P}(\bar{Z}_t \in A) := E[M_t I_{\{Z_t \in A\}}], \quad (57)$$

where $M_t := e^{-\int_0^t q(s, Z_s)ds}$. Moreover, for any Borel measurable and bounded function f ,

$$\begin{aligned} \bar{E}[f(\bar{Z}_t) | \bar{Z}_s = z] &= E[f(Z_t) \frac{M_t}{M_s} | Z_s = z] \\ &= E[f(Z_t) e^{-\int_s^t q(u, Z_u)du} | Z_s = z]. \end{aligned} \quad (58)$$

If Z is given by (51), letting

$$\mathcal{L}_t f(x) := \beta(t) \nabla f(x) + \frac{1}{2} \text{tr}(\delta(t) \delta^T(t) \nabla^2 f(x)) - q(t, x) f(x), \quad f \in \mathcal{C}_0^\infty(\mathbb{R}^d),$$

applying Itô's formula to $M_t f(Z_t)$ and taking expectation, we get

$$E[M_t f(Z_t)] = E[f(Z_0)] + \int_0^t E[M_s \mathcal{L}_s f(Z_s)] ds.$$

Thus, in view of (57), we have

$$\begin{aligned} \bar{E}[f(\bar{Z}_t)] &= \bar{E}[f(\bar{Z}_0)] + \int_0^t \bar{E}[\mathcal{L}_s f(\bar{Z}_s)] ds \\ &= \bar{E}[f(\bar{Z}_0)] + \int_0^t \bar{E}[\bar{E}[\mathcal{L}_s f(\bar{Z}_s) | \bar{Z}_s]] ds \\ &= \bar{E}[f(\bar{Z}_0)] + \int_0^t \bar{E}[\hat{\mathcal{A}}_s f(\bar{Z}_s)] ds, \end{aligned} \quad (59)$$

where,

$$\hat{\mathcal{A}}_t f(x) := \mathcal{A}_t f(x) - q(t, x) f(x) \quad (60)$$

and

$$\mathcal{A}_t f(x) := \bar{b}(t, x) \nabla f(x) + \frac{1}{2} \text{tr}(\bar{\sigma} \bar{\sigma}^T(t, x) \nabla^2 f(x)), \quad (61)$$

with

$$\begin{aligned} \bar{b}(t, x) &:= \bar{E}[\beta(t) | \bar{Z}_t = x] = E[M_t \beta(t) | Z_t = x], \\ \bar{\sigma} \bar{\sigma}^T(t, x) &:= \bar{E}[\delta(t) \delta^T(t) | \bar{Z}_t = x] = E[M_t \delta(t) \delta^T(t) | Z_t = x]. \end{aligned} \quad (62)$$

In view of Proposition 4, then there exist an \mathbb{R}^d -valued Brownian motion B , and a process $\hat{\bar{Z}}$ satisfying

$$\hat{\bar{Z}}_t = \hat{\bar{Z}}_0 + \int_0^t b(s, \hat{\bar{Z}}_s) ds + \int_0^t \sigma(s, \hat{\bar{Z}}_s) dB_s, \quad (63)$$

whose infinitesimal generator is $\hat{\mathcal{A}}$, such that, for each $t \geq 0$,

$$(\bar{Z}_t, \bar{E}[\beta(t) | \bar{Z}_t], \bar{E}[\delta(t) \delta^T(t) | \bar{Z}_t]) \stackrel{d}{=} (\hat{\bar{Z}}_t, b(t, \hat{\bar{Z}}_t), \sigma(t, \hat{\bar{Z}}_t) \sigma^T(t, \hat{\bar{Z}}_t)). \quad (64)$$

In terms of the mimicked killed Markov diffusion process $\hat{\bar{Z}}$, using (58) and (64), we have the following property for the t -reserve:

$$\begin{aligned} E[v^Z(t, Z_t)] &= E\left[\int_t^T e^{-\int_t^s r(u) du} E\left[e^{-\int_t^s q(u, Z_u) du} g(s, Z_s) | Z_t\right] ds\right] \\ &= \int_t^T e^{-\int_t^s r(u) du} \bar{E}[g(s, \bar{Z}_s)] ds \\ &= \int_t^T e^{-\int_t^s r(u) du} \bar{E}[g(s, \hat{\bar{Z}}_s)] ds \\ &= \int_t^T e^{-\int_t^s r(u) du} E\left[e^{-\int_t^s q(u, \hat{\bar{Z}}_u) du} g(s, \hat{\bar{Z}}_s)\right] ds \\ &= E[v^{\hat{\bar{Z}}}(t, \hat{\bar{Z}}_t)]. \end{aligned} \quad (65)$$

Applying the Feynman-Kac formula, $v = v^{\hat{\bar{Z}}}$ satisfies the following PDE

$$\begin{cases} \frac{\partial v}{\partial s}(s, x) + \hat{\mathcal{A}}_s v(s, x) + g(s, x) = r(s) v(s, x), & t \leq s < T, \\ v(T, x) = 0. \end{cases} \quad (66)$$

Hence, using (60), we get

$$\begin{cases} \frac{\partial v}{\partial s}(s, x) + \mathcal{A}_s v(s, x) + g(s, x) = (q(s, x) + r(s)) v(s, x), & t \leq s < T, \\ v(T, x) = 0. \end{cases} \quad (67)$$

Note that (65) does not imply that $v^Z(t, x) = v^{\widehat{Z}}(t, x)$ for all x , only that, 'on average over all x ' they will agree. A way to think of this is that if $v^Z(t, Z_t)$ is an unbiased estimator of some parameter θ , then $v^{\widehat{Z}}(t, \widehat{Z}_t)$ is also an unbiased estimator of θ . For all purposes, the PDE (67) is only useful if we can explicitly compute the terms \bar{b} and $\bar{\sigma}$ displayed in (62), which is in general out of reach even for the simplest Gaussian dynamics, due to presence of the path-dependent discounting factor M . This makes the idea of mimicking the killed process less attractive. We make one final attempt in constructing a mimicking process that preserves some properties of V_t .

4.2.3 Mimicking the environment process

We suggest the following recipe for computing an approximate t -reserve. First, we determine the Markovian projection \widehat{Z} of the underlying process Z . Then, we consider the moments of \widehat{V}_t defined by

$$v_n(t, z) = E^{t, z}[\widehat{V}_t^n], \quad (68)$$

which satisfies (14), as an approximation of the true moments based on Z . Using Proposition 4, letting

$$\begin{aligned} \alpha(t, z) &:= E[\beta(t)|Z_t = z] = E[\dot{a}(t)^T \nu_t + a(t)^T \mu | Z_t = z] \\ \gamma(t) &:= \sqrt{E[\delta(t)\delta(t)^T | Z_t = z]} = \sqrt{\delta(t)\delta(t)^T} = \sqrt{a(t)^T A A^T a(t)}, \end{aligned} \quad (69)$$

then the process \widehat{Z} defined by

$$d\widehat{Z}_t = \alpha(t, \widehat{Z}_t)dt + \gamma(t)d\widehat{W}_t, \quad (70)$$

where \widehat{W} is a standard Brownian motion, has the same marginal distributions as the process Z . However, this does not imply that V_t and \widehat{V}_t have the same marginals. In the numerical results section below, we study the distributions of V_t and \widehat{V}_t by Monte Carlo simulation of the processes ν and \widehat{Z} , respectively. It turns out that the distributions are almost identical, and we proceed with this mimicking approach. It then remains to determine the function α . We have

$$\begin{aligned} \alpha(t, z) &= E[\beta(t)|Z_t = z] = E[\dot{a}^T \nu_t + a^T \mu | a^T \nu_t = z] \\ &= a^T \mu + \dot{a}^T (\nu_0 + \mu t) \\ &\quad + E[\dot{a}^T A W_t | a^T A W_t = z - a^T (\nu_0 + \mu t)]. \end{aligned} \quad (71)$$

Since all linear combinations of W_t are Gaussian, we have

$$E[\dot{a}^T A W_t | a^T A W_t = z - a^T (\nu_0 + \mu t)] = (z - a^T (\nu_0 + \mu t)) \frac{\text{Cov}(\dot{a}^T A W_t, a^T A W_t)}{\text{Var}(a^T A W_t)}. \quad (72)$$

Using the independence of the marginal distributions of the components of W , we have

$$\begin{aligned} \text{Var}(a^T A W_t) &= \text{Var}\left(\sum_j W_t^j \sum_i a_i A_{ij}\right) \\ &= \sum_j \text{Var}(W_t^j) \left(\sum_i a_i A_{ij}\right)^2 = t a^T A A^T a = t a^T \Sigma a. \end{aligned} \quad (73)$$

Similarly,

$$\text{Cov}(\dot{a}^T AW_t, a^T AW_t) = ta^T \Sigma \dot{a}. \quad (74)$$

Finally, we obtain the following explicit expression for α ,

$$\alpha(t, z) = a^T \mu + \dot{a}^T (\nu_0 + \mu t) + (z - a^T (\nu_0 + \mu t)) \frac{a^T \Sigma \dot{a}}{a^T \Sigma a}. \quad (75)$$

Curiously, it happens that \hat{Z} is a Hull-White process, a model form which allows for explicit pricing of discount factors, see Hull and White [12]. Here, the hazard rate is given by the non-negative process $f(\hat{Z}_t)$, which is no longer of Hull-White form. Hence, we are unable to exploit the tractability of the Hull-White model. This is not necessarily a bad thing, since the Hull-White process allows for negative hazard rates, a property which is not always desired. Still, we may use Proposition 2 to compute moments of \hat{V}_t . From the representation (70), (14) becomes

$$\begin{cases} -\frac{\partial v_n}{\partial s} + n(f(z) + r(s))v_n = \alpha(s, z) \frac{\partial v_n}{\partial z} + \frac{1}{2} \gamma^2(s) \frac{\partial^2 v_n}{\partial z^2} + ng(s, z)v_{n-1}, & t \leq s < T \\ v_n(T, z) = 0, \end{cases} \quad (76)$$

with f , α and γ given by (32), (75) and (69), respectively. The PDE (76) can be solved using numerical methods, e.g. finite-difference schemes.

5 Numerical results

In this section, we implement two disability termination models together with the dimension reduction techniques of Sections 4.2.1 and 4.2.3. The parameters of the models for the years 2000-2011 are estimated using the method from [1]. Using Monte Carlo simulations, the distribution of the functional V_t is compared to the distributions of \hat{V}_t^1 and \hat{V}_t^2 , where \hat{V}_t^1 denotes the functional of the multivariate Markov process constructed in Section 4.2.1, and \hat{V}_t^2 denotes the functional of the Markov projection process of Section 4.2.3. Further, the PDE (76) is used to compute the first three moments of \hat{V}_t^2 , where we have chosen the parameters $x = 55$, $T = 10$, $r = 0.02$, $t = 0$, $g(t, z) = 1$. The PDE is solved using a first order implicit finite-difference scheme, and the results are compared to a Monte Carlo simulation with 100,000 draws and $\Delta t = 0.01$.

5.1 A linear model

We consider the model from Section 4 with basis functions

$$\begin{aligned} \phi^1(x) &= \frac{64 - x}{39}, & \phi^2(x) &= \frac{x - 25}{39} \\ \psi^1(t) &= 1, & \psi^2(t) &= t. \end{aligned}$$

We assume that ν follows a 4-dimensional Brownian motion, and estimate the drift and covariance matrix from the time series of parameter values.

The densities and distribution functions of V_t , \hat{V}_t^1 and \hat{V}_t^2 are presented in Figures 1-4. Note that, due to confidentiality, the x -axes are presented as fractions of the Best

Estimate anno 2011. Here, the Best Estimate is defined as the value of the initial reserve assuming that the model parameters are held constant over the entire policy period.

As can be seen in the plots, the density- and distribution functions of \widehat{V}_t^1 and \widehat{V}_t^2 are almost identical to those of V_t . Indeed, using a standard two-sample Kolmogorov-Smirnov test, we cannot reject the hypothesis that the samples of V_t and \widehat{V}_t^1 are drawn from the same distribution. The corresponding p -value is 0.56. However, we can in fact reject the hypothesis that the samples of V_t and \widehat{V}_t^2 are drawn from the same distribution. Still, we conclude that we can consider \widehat{V}_t^2 as an approximation of V_t , and that, as expected, V_t and \widehat{V}_t^1 have identical distributions. This is a highly useful result since it reduces the dimensionality of the problem, which significantly reduces the computational cost. In this example, the choice stands between obtaining an exact result with two space dimensions, or an approximate result with one space dimension, compared to the four space dimensions of the original problem.

Numerical values from the PDE solver for the first three moments, as a fraction of the Best Estimate anno 2011, are presented in Table 1. The values of v_1 correspond to the initial reserve. We present the values as fractions of the Best Estimate rather than monetary units due to confidentiality. 99% approximate confidence intervals of moments of V_t , \widehat{V}_t^1 and \widehat{V}_t^2 from the Monte Carlo simulation are presented in Table 2. As we can see, the moments from the PDE solver lie well within the 99% confidence intervals from the Monte Carlo simulation of \widehat{V}_t^2 , and a few percentage points above the 99% confidence intervals from the Monte Carlo simulation of V_t . We stress the fact that we are trading accuracy for computational efficiency.

Table 1: Moments of \widehat{V}_t^2 from the PDE solver, scaled by the Best Estimate.

$\Delta z = \Delta t$	v_1	v_2	v_3
0.1	0.9097	0.8019	0.7567
0.05	0.9064	0.7929	0.7370
0.01	0.9040	0.7865	0.7239
0.005	0.9037	0.7858	0.7226
0.001	0.9035	0.7853	0.7217

Table 2: 99% approximate confidence intervals of moments of V_t , \widehat{V}_t^1 and \widehat{V}_t^2 from the Monte Carlo simulation.

	v_1	v_2	v_3
MC, V_t	(0.8986 0.9011)	(0.7720 0.7772)	(0.6938 0.7030)
MC, \widehat{V}_t^1	(0.8991 0.9016)	(0.7726 0.7778)	(0.6944 0.7035)
MC, \widehat{V}_t^2	(0.9013 0.9041)	(0.7812 0.7870)	(0.7148 0.7257)

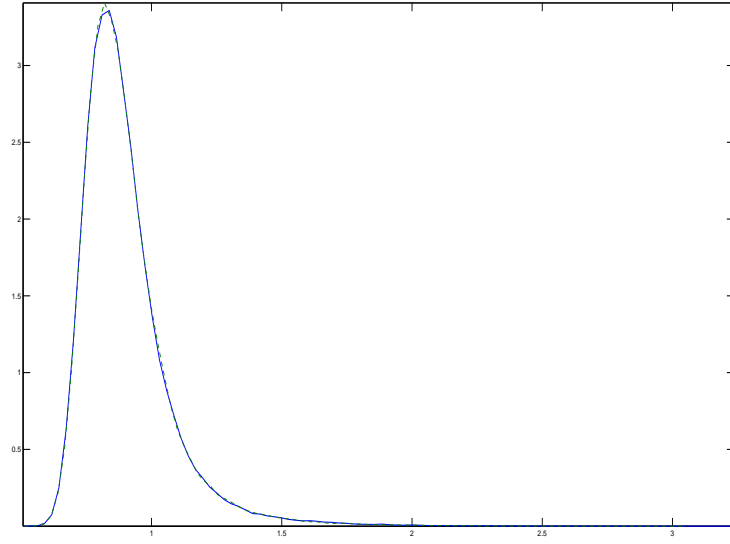


Figure 1: The densities of V_t (solid) and \widehat{V}_t^1 (dashed).

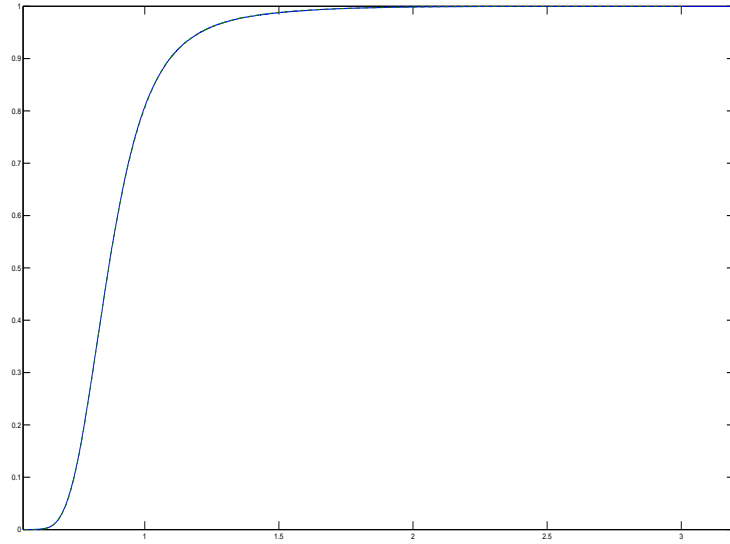


Figure 2: The distribution functions of V_t (solid) and \widehat{V}_t^1 (dashed).

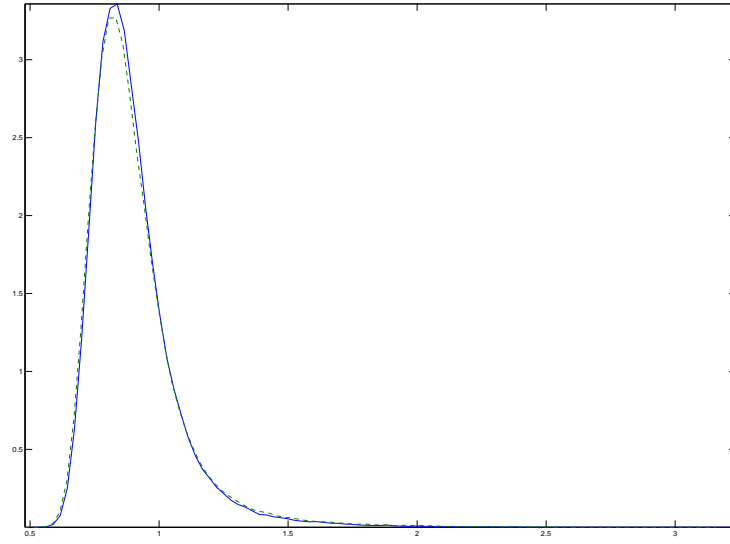


Figure 3: The densities of V_t (solid) and \hat{V}_t^2 (dashed).

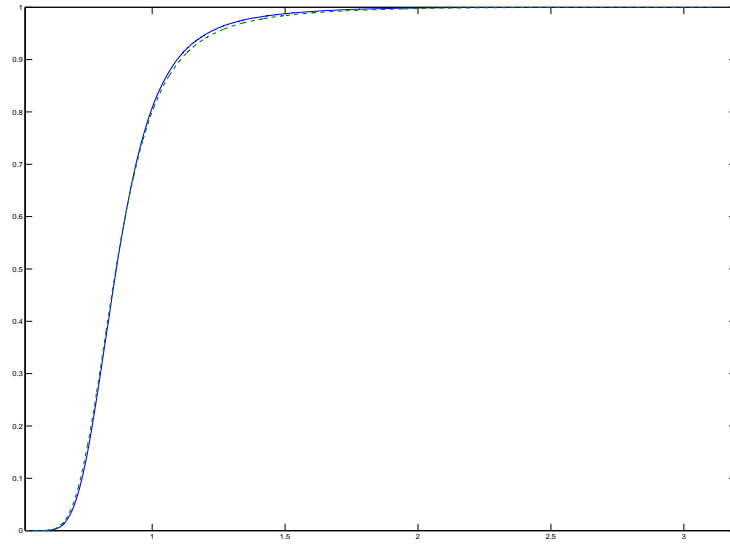


Figure 4: The distribution functions of V_t (solid) and \hat{V}_t^2 (dashed).

5.2 A non-linear model

Next, we consider the model from Section 4 with basis functions

$$\begin{aligned}\phi^1(x) &= \frac{64-x}{39} \quad , \quad \phi^2(x) = \frac{x-25}{39} \\ \psi^1(t) &= 1 \quad , \quad \psi^2(t) = e^{-t} \quad , \quad \psi^3(t) = e^{-2t}.\end{aligned}$$

Using the method from [1], this model yields slightly better goodness of fit compared to the linear model. We assume that ν follows a 6-dimensional Brownian motion, and estimate the drift and covariance matrix from the time series of parameter values.

For this non-linear model, it is immediate that we cannot implement the recipe of Section 4.2.1 to reduce the dimensionality. Instead, we focus our efforts on the Markov projection technique of Section 4.2.3.

The densities and distribution functions of V_t and \widehat{V}_t^2 are presented in Figures 5-6, and 99% approximate confidence intervals of moments of V_t and \widehat{V}_t^2 from the Monte Carlo simulation are presented in Table 3. It is apparent that the mimicking approximation performs slightly worse for this model compared to the linear model. However, comparing Table 3 with Table 2, it seems that the approximation error and the model uncertainty are of the same magnitude: the deviations between the linear model and the non-linear model are similar to the deviations between any one of the models and its corresponding Markovian projection, at least for the first two moments. For the non-linear model, the Markovian projection shows a significant approximation error for the third moment. Again, we stress the fact that we are trading accuracy for computational efficiency. As the Markov projection technique seems to slightly overestimate both the moments and the thickness of the tail of V_t , it could possibly be used to obtain conservative risk estimates, although further research is needed to confirm this hypothesis.

Table 3: 99% approximate confidence intervals of moments of V_t and \widehat{V}_t^2 from the Monte Carlo simulation.

	v_1	v_2	v_3
V_t	(0.9292 0.9313)	(0.8156 0.8200)	(0.7368 0.7441)
\widehat{V}_t^2	(0.9381 0.9415)	(0.8605 0.8688)	(0.8583 0.8768)

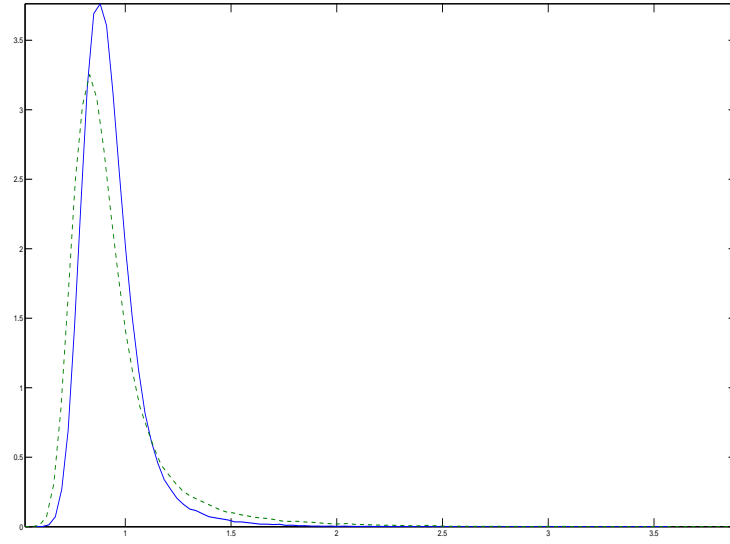


Figure 5: The densities of V_t (solid) and \widehat{V}_t^2 (dashed).

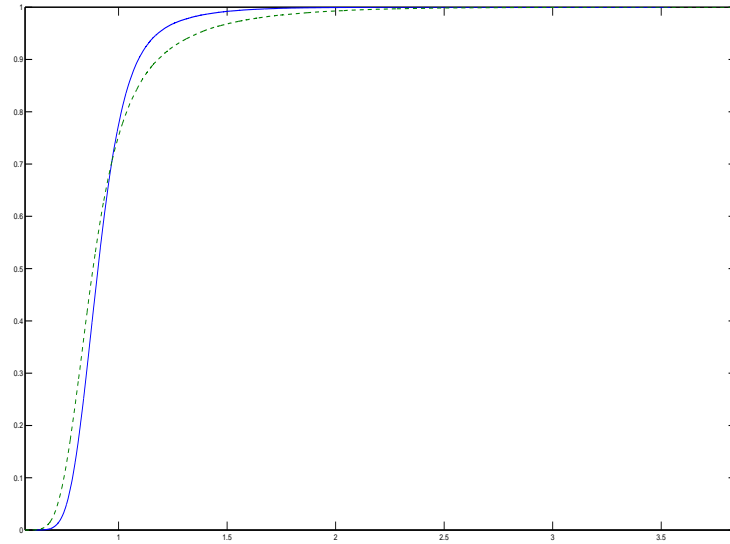


Figure 6: The distribution functions of V_t (solid) and \widehat{V}_t^2 (dashed).

6 Acknowledgements

The first author gratefully acknowledges financial support from the Swedish Export Credit Corp. (SEK). The second author gratefully acknowledges financial support from the Filip Lundberg and Eir's 50 Years foundations. Both authors appreciate the helpful comments of an anonymous referee.

References

- [1] ARO, H., DJEHICHE, B., AND LÖFDAHL, B. Stochastic modelling of disability insurance in a multi-period framework. *Scandinavian Actuarial Journal* (2013).
- [2] BENTATA, A., AND CONT, R. Mimicking the marginal distributions of a semi-martingale. *arXiv:0910.3992* (2012).
- [3] BHATT, A., AND BORKAR, V. Occupation measures for controlled markov processes: Characterization and optimality. *Annals of Probability* 24 (1996), 1531–1562.
- [4] BIFFIS, E. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics* 37 (2005), 443–468.
- [5] BORKAR, V. Mimicking finite dimensional marginals of a controlled diffusion by simpler controls. *Stochastic Processes and their Applications* 31, 2 (1989), 333–342.
- [6] BOUHADOU, S., AND OUKNINE, Y. Mimicking finite dimensional marginals of a controlled diffusion with jumps. *Stochastics and Dynamics* 14, 1 (2014).
- [7] CHRISTIANSEN, M., DENUIT, M., AND LAZAR, D. The Solvency II square-root formula for systematic biometric risk. *Insurance: Mathematics and Economics* 50 (2012), 257–265.
- [8] DAHL, M. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics* 35 (2004), 113–136.
- [9] DAHL, M., AND MØLLER, T. Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics and Economics* 39 (2006), 193–217.
- [10] FRIEDMAN, A. *Stochastic differential equations and applications*. Courier Dover Publications, 2012.
- [11] GYÖNGY, I. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability Theory and Related Fields* 71, 4 (1986), 501–516.
- [12] HULL, J., AND WHITE, A. Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies* 3, 4 (1990), 573–592.
- [13] HYNDMAN, R., AND ULLAH, M. S. Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis* 51 (2007), 4942–4956.
- [14] KRYLOV, N. Once more about the connection between elliptic operators and Itô’s stochastic equations. In *Statistics and control of stochastic processes, Steklov Seminar* (1984), pp. 214–229.
- [15] KURTZ, T. G., AND STOCKBRIDGE, R. H. Existence of markov controls and characterization of optimal markov controls. *SIAM Journal on Control and Optimization* 36 (1998), 609–653.

- [16] KURTZ, T. G., AND STOCKBRIDGE, R. H. Stationary solutions and forward equations for controlled and singular martingale problems. *Electronic Journal of Probability* 6, 17 (2001), 1–52.
- [17] LEVANTESI, S., AND MENZIETTI, M. Managing longevity and disability risks in life annuities with long term care. *Insurance: Mathematics and Economics* 50 (2012), 391–401.
- [18] LUDKOVSKI, M., AND YOUNG, V. R. Indifference pricing of pure endowments and life annuities under stochastic hazard and interest rates. *Insurance: Mathematics and Economics* 42 (2008), 14–30.
- [19] NORBERG, R. Differential equations for moments of present values in life insurance. *Insurance: Mathematics and Economics* 17 (1995), 171–180.
- [20] NORBERG, R. Forward mortality and other vital rates - Are they the way forward? *Insurance: Mathematics and Economics* 47 (2010), 105–112.
- [21] ØKSENDAL, B. *Stochastic differential equations*. Springer, 2003.
- [22] PRAKASA RAO, B. L. S. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics* 61, 2 (2009), 441–460.
- [23] ROGERS, L., AND WILLIAMS, D. *Diffusions, Markov processes and martingales. Volume one: Foundations*. Wiley, 1995.
- [24] VASICEK, O. The distribution of loan portfolio value. *Risk* 15, 12 (2002), 160–162.